Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 66 (2013) 425 – 434

Procedia Engineering

www.elsevier.com/locate/procedia

5th Fatigue Design Conference, Fatigue Design 2013

Detection of the Curves based on Lateral Acceleration using Hidden Markov Models

Roza Maghsood[a]*, Pär Johannesson[b]

[a]*Chalmers University of Technology , SE-412 96 Gothenburg, Sweden*
[b]*SP Technical Research Institute of Sweden, SE-400 22 Gothenburg, Sweden*

Abstract

In vehicle design it is desirable to model the loads by describing the load environment, the customer usage and the vehicle dynamics. In this study a method will be proposed for detection of curves using a lateral acceleration signal. The method is based on hidden Markov models (HMMs) which are probabilistic models that can be used to recognize patterns in time series data. In an HMM, 'hidden' refers to a Markov chain where the states are not observable, however what can be observed is a sequence of data where each observation is a random variable whose distribution depends on the current hidden state. The idea here is to consider the current driving event as the hidden state and the lateral acceleration signal as the observed sequence. Examples of curve detection are presented for both simulated and measured data. The classification results indicate that the method can recognize left and right turns with small misclassification errors.

Keywords: Hidden Markov models (HMMs); Markov chain; curve detection; event classification; lateral acceleration.

1. Introduction

For fatigue design the loads need to be assessed. One approach is to describe the load environment and the customer usage, which together with the vehicle dynamics define the load conditions. The characteristics of driving events used for describing customer usage can be defined using measurements obtained from specially equipped vehicles on a test track. On the other hand, measuring on vehicles in service is difficult and expensive and in general

* Corresponding author. Tel.: +46 31 772 53 79.
E-mail address: rozam@chalmers.se

1877-7058 © 2013 The Authors. Published by Elsevier Ltd.
Selection and peer-review under responsibility of CETIM
doi:10.1016/j.proeng.2013.12.096

there is no access to measurements dedicated to durability. Thus, for on-board logging of events we need to use the information which is available for all vehicles by means of CAN (Controller Area Network) bus data.

A data set is available from Volvo Trucks, and the important events have been defined based on dedicated test track measurements. The problem is to identify the events from CAN-data and their frequencies. In this study we propose a method using hidden Markov models (HMMs) to detect curves based on a lateral acceleration signal. The idea is that we can consider the driving events, i.e. straights and curves, as the hidden states and construct the model based on them.

The HMMs have been widely used in signal processing to recognize the events and also to predict them in the future, see e.g. the overview by Rabiner [8] with applications to speech recognition. Mitrović [6, 7] and Berndt and Dietmayer [2] used HMMs to detect driving events. They constructed one HMM for each type of driving event such as left and right curves, left, right and straight on roundabout. They created a training set by identifying events manually to build the models and evaluate them. Then for a new observation sequence, they computed the observation likelihoods based on all models and chose the driving event type with respect to the highest likelihoods.

The parameters in an HMM are the transition probability matrix, the emission matrix and the initial state distribution. They must be estimated to characterize the model. In our suggested method, we have used a single HMM for describing all events instead of constructing several different models where each HMM describes a single event. It should be simpler to estimate the parameters of one model than lots of parameters of different models.

In an HMM, a training set is used to estimate the parameters of the model, while a test set is used to validate the model. A training set consists of all necessary information for estimating the model parameters. In our study, the training set contains all history about the curves such as the start and stop points of them. We have simulated different lateral acceleration signals with different lengths and different number of curves (events) to have some controlled training and test sets.

In Section 2 we describe the concept of HMMs and present two methods for detection of curves. For method 1 the parameters of the HMM is estimated from the training set, while for method 2 the transition matrix is re-estimated based on the test set. Examples and their results for simulated and measured data are shown in Section 3. Conclusions are presented in Section 4.

2. Hidden Markov models

Hidden Markov models are probabilistic models that can be used for detection of patterns or events in a signal. The setup is that there are two processes. The interesting process Z_t that describes the events is not possible to measure. It is thus called hidden and modelled as a Markov chain. However, what can be observed is a process Y_t whose statistical properties depend on the value of Z_t. The problem at hand is to estimate the parameters of the HMM. Based on an observation of Y_t, it is then possible to reconstruct the most probable hidden process and identify events.

In this study, three events right turn (RT), left turn (LT) and straight forward (SF) have been considered. The idea is that one can see these three events as three hidden states and construct the HMM based on them. Fig. 1 illustrates three hidden states and a sequence of observations that can be generated based on the probability distribution of observation symbols.

Let $\{Z_t\}_{t=1}^{\infty}$ be a Markov chain where Z_t denotes a hidden state at time t and has possible values $S = \{S_1, S_2, \ldots, S_N\}$. The transition probabilities between the hidden states are defined by the matrix $A = \{a_{ij}\}$, called transition matrix, where

$$a_{ij} = P(Z_{t+1} = S_j | Z_t = S_i) \qquad (1)$$

for $i, j = 1,2,..., N$ and $\sum_{j=1}^{N} a_{ij} = 1$.

Further, there is another process $\{Y_t\}_{t=1}^{\infty}$ denoting the observation symbol at time t. The sequence of observations

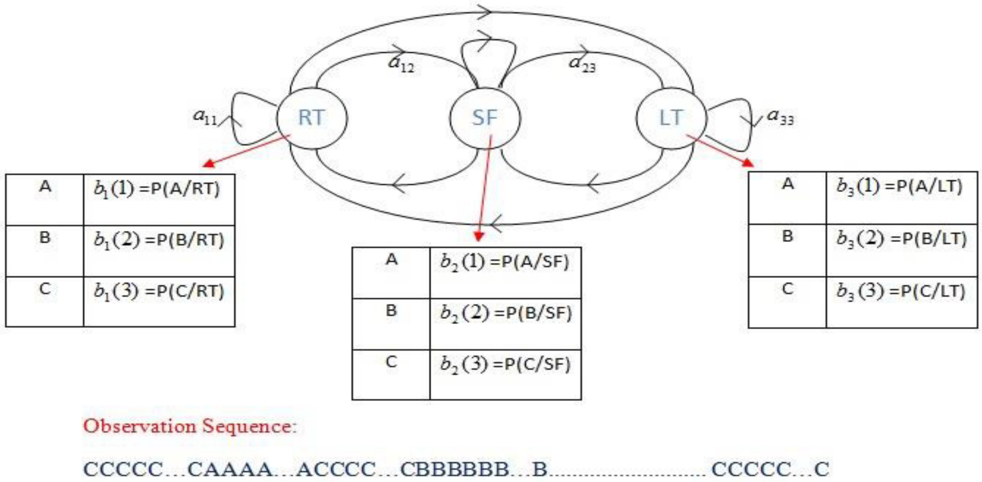

Fig. 1. The hidden state sequence is modeled by a Markov chain and the observation sequence is modeled by the emission probabilities.

has possible values $V = \{V_1, V_2, \ldots, V_M\}$ and it is observable for us. The probability distribution of observation symbols in each state is given by the emission matrix, $B = \{b_j(V_k)\}$, where

$$b_j(V_k) = P(Y_t = V_k | Z_t = S_j) \tag{2}$$

and $\sum_{j=1}^{N} b_j(V_k) = 1$.

The state where the hidden process will start is modeled by the initial state probabilities that are denoted by $\pi = \{\pi_1, \pi_2, \ldots, \pi_N\}$ where

$$\pi_i = P(Z_1 = S_i) \tag{3}$$

for $i = 1,2,...,N$ and $\sum_{i=1}^{N} \pi_i = 1$.

It has been demonstrated that a discrete HMM can be good in pattern recognition, see Rabiner [8]. We have also used a discrete HMM $\lambda = (A, B, \pi)$ where λ represents model parameters which contain the transition matrix, the emission matrix and the initial state distribution. .

As mentioned before, we have three hidden states $S = \{\mathrm{RT}, \mathrm{SF}, \mathrm{LT}\}$ denoting the three events right turn, straight forward and left turn, respectively. In order to estimate the parameters of the HMM, we have used the lateral acceleration signal where we also have an observation of the hidden process Z_t. This will be our training data that contains observation of both the Y-process and the hidden Z-process. We have considered lateral acceleration values as our data and thus we need to translate this continuous feature into predefined classes. Here, three classes will be used, $V = \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, that are defined as follows:

$$\begin{aligned} \mathrm{A} &= \{\text{"lateral acceleration"} < -0.2\ \mathrm{m/s^2}\}, \\ \mathrm{B} &= \{-0.2\ \mathrm{m/s^2} \le \text{"lateral acceleration"} \le 0.2\ \mathrm{m/s^2}\}, \\ \mathrm{C} &= \{\text{"lateral acceleration"} > 0.2\ \mathrm{m/s^2}\}. \end{aligned} \tag{4}$$

This kind of clustering will create a sequence of observation symbols which has been used to estimate the emission matrix in our model.

To estimate the transition probabilities, we have just counted the number of transitions between the three states and normalized each row of the transition matrix to one. To estimate the emission matrix, we have counted the number of times that each observation symbol has been seen in each state.

2.1. Model Evaluation

The aim of this study is to find a probabilistic model to recognize the curves. We have estimated the parameters of the model by using a training set and evaluated it by using different new sequences of observations as our test set. To identify the curves for a new lateral acceleration signal, we have considered two different methods as following:

- **Method 1:** Use the estimated transition and emission matrices from the training set.
- **Method 2:** Use the emission matrix from the training set but re-estimate the transition matrix based on the new signal.

The main reason for considering method 2 is the differences between roads that can affect on the transitions between states. The emission matrix describes the property of the curves given certain hidden states, however, the transition matrix describes the duration of the events. Thus, it could be reasonable to update the transition matrix for a new signal to find the hidden states.

2.1.1. Method 1

Here, we have used a training set to estimate the parameters $\lambda = (A, B, \pi)$ of the HMM. The Viterbi algorithm, see Viterbi [9] and Forney [4], has been used to find the most probable sequence of hidden states for a new signal.

Suppose that we have classified the new lateral acceleration values with length n and we got an observation sequence $y_1, y_2, ..., y_n$. We would like to find driving events for this new observation. It means that we want to find a sequence of hidden states which maximizes the probability of observing this specified observation. The Viterbi algorithm finds the state sequence $z_1, z_2, ..., z_n$ out of the 3^n possible sequences of length n that maximizes

$$P\left(Y_1 = y_1, ..., Y_n = y_n \middle| Z_1 = z_1, ..., Z_n = z_n; \lambda\right). \tag{5}$$

In fact, the Viterbi algorithm gives a state sequence $z_1, z_2, ..., z_n$ that maximizes the conditional probability of the observation sequence for given parameters $\lambda = (A, B, \pi)$. The result will give the most likely sequence of hidden states from which it is possible to identify the driving events.

2.1.2. Method 2

In this approach, we have fixed the emission matrix from the training set and re-estimated the transition matrix from each new signal. To estimate model parameters based on an observation sequence, we have used the Baum-Welch algorithm which was introduced by Baum et al. [1]. It is equivalent to the EM (expectation-maximization) algorithm, see Dempster et al. [3]. The Baum-Welch algorithm is one of the well known methods to estimate the model parameters in HMMs. It is an iterative maximum likelihood method and starts with initial parameters that in our case are set based on training data. The algorithm uses a forward-backward procedure to estimate the model parameters for a given sequence of observations. See Rabiner [8] where the general EM algorithm in HMMs is described.

In method 2, we didn't update the probabilities $b_j(V_k)$ in the Baum-Welch algorithm since we have fixed the emission matrix $B = B_{\text{training}}$. We have just re-estimated the transition matrix and the initial state distribution for a given observation sequence. Then, we have used Viterbi algorithm to find the most likely hidden states based on given parameters.

3. Examples

We have tested our models with simulated and measured data sets. In the first example, we have simulated different lateral acceleration signals as our training and test sets and evaluated the detection of driving events. In the

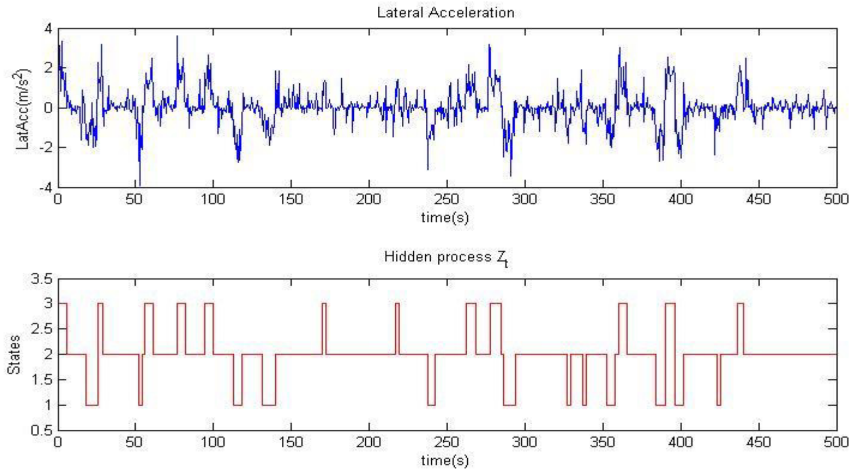

Fig. 2. Simulated lateral acceleration signal and the corresponding hidden states.

second example, we have used measured data which is dedicated field measurements from a Volvo Truck.

3.1. Simulated lateral acceleration signal

We need a training set to estimate the parameters and a test set to evaluate the model. For this purpose, we have simulated different lateral acceleration signals. Fig. 2 shows an example of the simulated lateral acceleration signal and the corresponding hidden states. These two simulated signals will be our training set. Next, we will describe how the simulation has been performed.

At first we have generated the events by using a Markov chain in our simulation. We supposed that the probabilities of going from a right turn to a left turn and vice versa are quite small. Most often we will have straight forward after a right turn or a left turn. Thus, we have considered a transition matrix such as:

$$P = \begin{bmatrix} 0 & 0.9 & 0.1 \\ 0.5 & 0 & 0.5 \\ 0.1 & 0.9 & 0 \end{bmatrix}, \tag{6}$$

and simulated a Markov chain with three states which represent our sequence of events.

Since we are going to model the length of each straight and each curve, we have chosen the start and stop points of each $event(i), i = 1,2,...,k$ as follows:

$$\text{Start point for } event(i) = \text{Stop point for } event(i-1),$$
$$\text{Stop point for } event(i) = \text{Start point for } event(i) + L_i. \tag{7}$$

where $\text{Start point for } event(1) = 0$. The length (duration) of each event L_i is random according to specified distributions, namely

- If $event(i)$ is a curve (right or left turn), then $L_i \sim U(2,8)$ since each turn may take between 2-8 seconds.
- If $event(i)$ is straight, then $L_i \sim Exp(\theta)$ where $\theta = 20$ shows the average duration of each straight.

The result will be our simulated hidden process Z_t. To generate a lateral acceleration signal, we have used a model suggested by Karlsson [5]. The measured lateral acceleration can be split into two load processes which are the centripetal acceleration and a residual. To get the Y_t process, we have translated lateral acceleration values into the symbols $V = \{A, B, C\}$ as described in Eq.(4).

3.1.1. Estimate of parameters from training set

Recall that in this example the signal in Fig. 2 will be our training set. The signal contains 500 events and we have considered the value $\theta = 20$ to get the duration of each straight. Fig. 2 illustrates only the first 1000 time points of the training set.

To estimate the transition matrix, we have counted the number of transitions between the three states. Finally, we have counted the number of times that each observation symbol A, B and C has been seen in each state to estimate the emission matrix. The transition matrix is:

$$A = \begin{bmatrix} 0.901 & 0.083 & 0.016 \\ 0.013 & 0.976 & 0.011 \\ 0.010 & 0.092 & 0.898 \end{bmatrix}. \tag{8}$$

The emission matrix is:

$$B = \begin{bmatrix} 0.965 & 0.020 & 0.015 \\ 0.182 & 0.636 & 0.182 \\ 0.011 & 0.023 & 0.966 \end{bmatrix}. \tag{9}$$

3.1.2. Model evaluation

To recognize the curves for a new simulated lateral acceleration signal, we have considered two different methods. We have generated a new lateral acceleration signal as our testing set to compare the two methods. The new signal is shorter than the training set and we have changed the value $\theta = 5 \times 20$ which means that we get long straights. The simulation contains 28 curves.

Method 1: Here, we have estimated both transition and emission matrices from the training set. Then, the Viterbi algorithm has been used to find the most probable sequence for the new signal. Fig. 3 shows the true and detected states based on our model. It can be seen that, for this signal, the method can recognize left turn and right turn with small misclassification error.

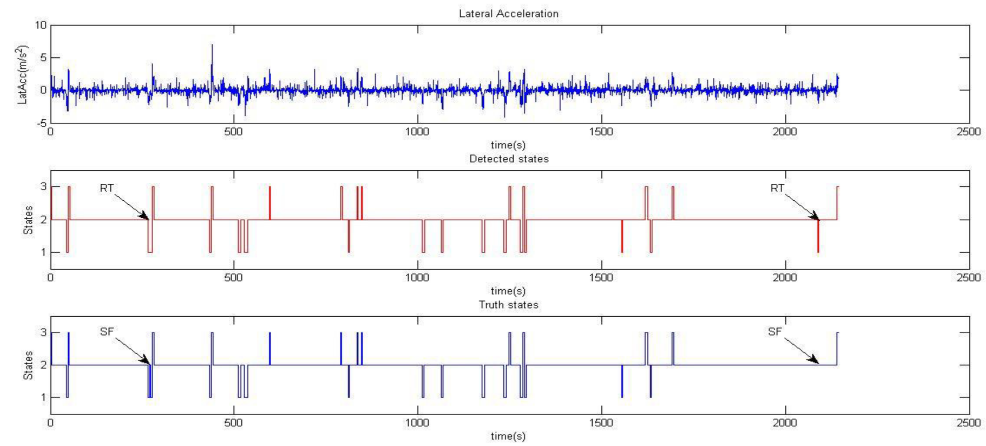

Fig. 3. Detection of events using method 1.

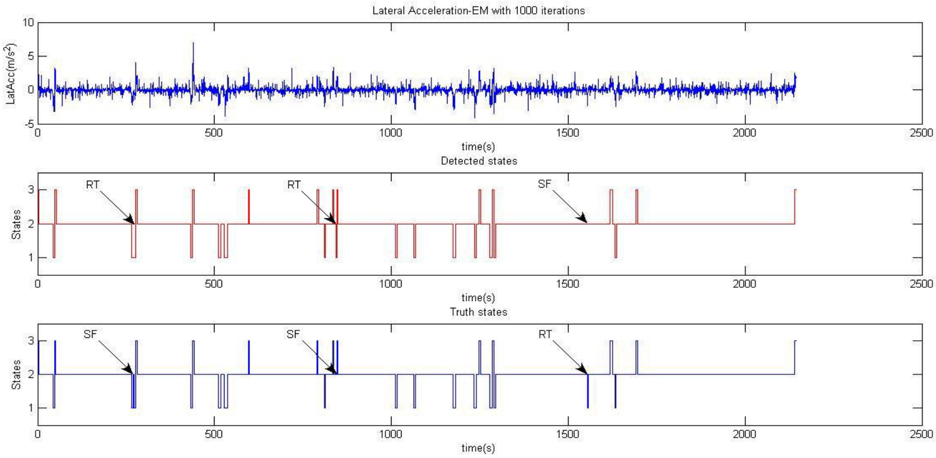

Fig. 4. Detection of events using method 2.

Method 2: Here, we have used the estimated emission matrix from the training set, but we have estimated the transition matrix from the new signal based on the EM algorithm. The re-estimated transition matrix is:

$$A = \begin{bmatrix} 0.880 & 0.088 & 0.032 \\ 0.005 & 0.992 & 0.003 \\ 0.014 & 0.134 & 0.852 \end{bmatrix}. \tag{10}$$

The true and detected states for the new signal are shown in Fig. 4, where we can see that the misclassification error rate in this case is higher than for method 1.

Comparison between method 1 and 2: To get the misclassification error rates, we have calculated both type I (false positive) and type II (false negative) errors. If we find an event that does not exist, we get a false positive error. However, if we can't detect the true event, the false negative error will happen.

Most of the time, the duration of the detected events are not the same as the real events. Therefore, we have considered the middle time of each detected event and we have compared its label with the true label (hidden state) at that time. The number of times that we got different labels divided by the number of events will be the false positive error rate. Further, to get the false negative error, we have considered the true label of each event at the middle and we have compared it with the detected label.

We did 1000 simulations to get an average of error rates. At first we have simulated 1000 signals with 100 events as our test sets, where the parameters of the model are the same as the ones in the training set. Therefore, $\theta = 20$ and the transition probabilities for Markov chain is:

$$P = \begin{bmatrix} 0 & 0.9 & 0.1 \\ 0.5 & 0 & 0.5 \\ 0.1 & 0.9 & 0 \end{bmatrix}. \tag{11}$$

The average false positive error based on method 1 is 0.025 and the average false negative error is 0.046.

For method 2, the average false positive error is 0.048 and the average false negative error is 0.034. The results are summarized in Table 1.

Table 1. Type I and Type II errors, where the training and test sets have the same parameters.

Error	Type I	Type II
Method 1	0.025	0.046
Method 2	0.048	0.034

Finally, we have changed the parameters for the test sets to check how much they will affect on the results. We have considered long straights by setting $\theta = 5 \times 20$. The modified transition matrix for the Markov chain is:

$$P = \begin{bmatrix} 0 & 0.85 & 0.15 \\ 0.6 & 0 & 0.4 \\ 0.05 & 0.95 & 0 \end{bmatrix}. \tag{12}$$

Table 2 shows the error rates regarding to the new parameters, where it can be seen that the difference between method 1 and 2 is small. Compared to Table 1, for method 1, the false positive error increases while the false negative decreases, whereas for method 2 the result is almost the same

Table 2. Type I and Type II errors, where the training and test sets have different parameters.

Error	Type I	Type II
Method 1	0.051	0.031
Method 2	0.047	0.038

If the parameters of the test set is similar to the ones in the training set, then method 1 should be preferred. The emission matrix is expected to be similar for all road types. However, the transition matrix should depend on the type of the road. This motivates the use of method 2. For example if we have a lateral acceleration signal from a city road as our training set and we want to detect events based on a lateral acceleration signal from a highway, then the transition matrix from the training set can not be good and it could be re-estimated from the new signal.

The simulation study indicates that method 1 is still quite robust to changes in the transition matrix, since it detects the events equally accurate as method 2, even though the transition matrix in training and test sets are different.

3.2. Measured lateral acceleration signal

The measured data that we have used is a field measurement coming from a Volvo Truck. We have used measured signals from the CAN bus and we have manually detected the events by looking at video recordings from the truck cabin to see what had happened during the driving. By having the start and stop points of each event, we have created the hidden Z-process. For the Y-process, we need a lateral acceleration signal which we have computed by using the following formula:

$$\textit{lateral Acceleration} = (\textit{Speed} \times \textit{Yaw Rate})/3.6. \tag{13}$$

To remove the high frequency noise, we have used a Butterworth low-pass filter with 0.5 Hz cut-off frequency. To reduce the amount of data, we have split the data into frames (the duration of each frame is 0.5 sec) and calculated the mean value for each frame. We have translated the continuous feature (mean value) into the predefined symbols in each frame by three classes A, B and C where

$$\begin{aligned} \mathrm{A} &= \{\text{"lateral acceleration"} < -0.5\ \mathrm{m/s^2}\}, \\ \mathrm{B} &= \{-0.5\ \mathrm{m/s^2} \le \text{"lateral acceleration"} \le 0.5\ \mathrm{m/s^2}\}, \\ \mathrm{C} &= \{\text{"lateral acceleration"} > 0.5\ \mathrm{m/s^2}\}. \end{aligned} \tag{14}$$

Compared to Eq. (4), we have changed the threshold from 0.2 to 0.5 in our clustering in order to improve the detection results.

The signal that is considered has the length 3800 seconds, which we have divided into two parts as our training and test sets. The training set contains 2000 seconds and the test set contains 1800 seconds. Fig. 5 shows the training part of the signal and the corresponding manually detected hidden states.

Method 1: At first, we have used method 1 and estimated transition and emission matrices from the training set, resulting in the transition matrix:

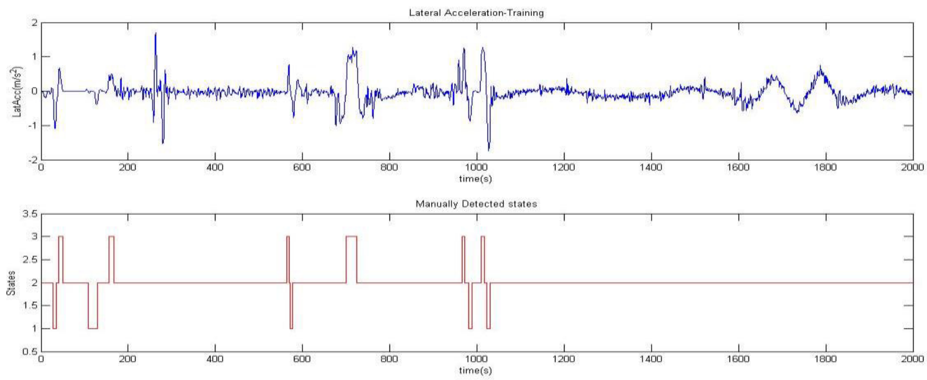

Fig. 5. Training part of measured lateral acceleration signal and the corresponding manually detected hidden states.

$$A = \begin{bmatrix} 0.945 & 0.055 & 0.000 \\ 0.001 & 0.997 & 0.002 \\ 0.000 & 0.048 & 0.952 \end{bmatrix}, \tag{15}$$

and the emission matrix is:

$$B = \begin{bmatrix} 0.418 & 0.582 & 0.000 \\ 0.031 & 0.957 & 0.012 \\ 0.000 & 0.363 & 0.637 \end{bmatrix}. \tag{16}$$

The detected states based on method 1 for the test set is shown in Fig. 6, where we can compare them with the manually detected states. It can be seen that the misclassification error rate is high. In all cases, the method can recognize the manually detected curves. However, we have a false positive error since the method has found five right turns that are not in the manual detection. One reason could be that the manual detections are not completely correct because of the visual errors and the low quality of videos. There is also a sharp left turn which could not be

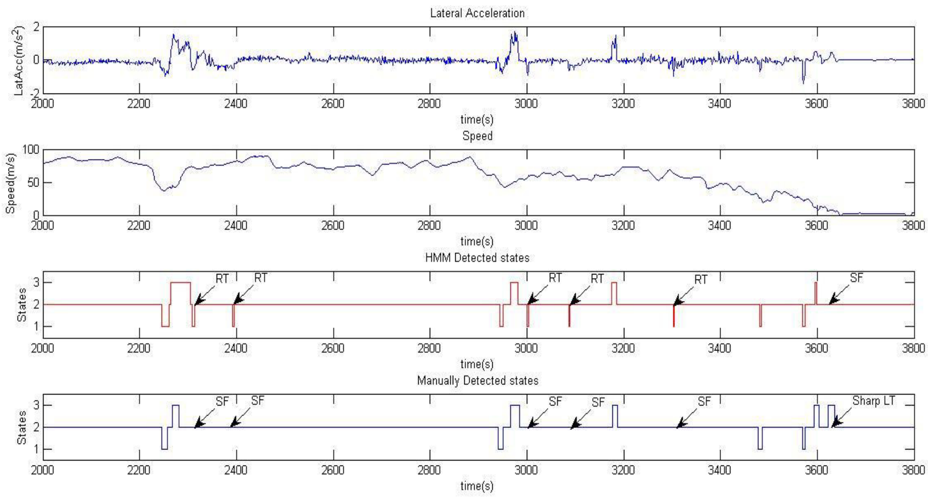

Fig. 6. Detection of events in the measured signal using method 1.

detected since the speed is low (about 10 km/s) at that part which makes it hard to recognize the curve correctly.

Method 2: In method 2, the transition matrix has been re-estimated based on the EM algorithm resulting in the estimated transition matrix:

$$A = \begin{bmatrix} 0.895 & 0.105 & 0.000 \\ 0.004 & 0.995 & 0.001 \\ 0.000 & 0.033 & 0.967 \end{bmatrix}. \tag{17}$$

The results of method 2 are the same as for method 1. Since the road type of the test set is similar to the one in the training set, method 1 and 2 performed similarly to detect the events.

4. Conclusion

The examples in this study indicate that the HMMs can be used to recognize the curves based on a lateral acceleration signal. We have considered three driving events (right turn, left turn and straight forward) as the hidden states and constructed the model based on them. The parameters of the model have been estimated by considering two different methods. In method 1, we have estimated the transition and emission matrices from the training set, while in method 2 the emission matrix has been fixed from the training set and the transition matrix has been re-estimated based on the test set.

The results of the simulation study show that method 1 should be preferred if the parameters of the test set is similar to the ones in the training set, i.e. the characteristics of the roads are likely to be similar. The emission matrix is expected to be similar for all road types, however, the transition matrix should depend on the type of the road. This motivates the use of method 2. If we have a lateral acceleration signal from a city road as our training set and we want to detect events based on a lateral acceleration signal from a highway, then the transition matrix from the training set can not be good and it could be re-estimated from the new signal.

The method can be extended to detect more events, such as braking and static steering, by considering more signals which contain useful information about the events. One approach can be to combine the essential signals and increase the number of classes in Y -process. It could then be possible to detect more events.

Acknowledgements

We are thankful to Dr. Thomas Svensson for his useful ideas and helpful suggestions in this study. We would like to thank Volvo Trucks for supplying the data in this study and to the members in our research group at Volvo for their valuable advice. Finally, we gratefully acknowledge the financial support from VINNOVA.

References

[1] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, *Ann. Math. Statist.,* 41(1), 1970, pp. 164- 171.

[2] H. Berndt and K. Dietmayer, Driver Intention Inference with Vehicle Onboard Sensors, *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, Nov. 11-12, Pune, 2009, pp. 102 - 107.

[3] A. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via EM algorithm, *J. Roy. Stat. Soc.,* 39(1), 1977, pp. 1-38.

[4] G. D. Forney, The Viterbi algorithm, *Proc. IEEE,* 61, 1973, pp. 268 - 278.

[5] M. Karlsson, *Load Modelling for Fatigue Assessment of Vehicles-a Statistical Approach*, PhD thesis, Chalmers University of Technology, 2007.

[6] D. Mitrović, *Learning Driving Patterns to Support Navigation*, PhD thesis, University of Canterbury, 2004.

[7] D. Mitrović, Reliable Method for Driving Events Recognition, *IEEE Trans. Syst*., 6(2), 2005, pp. 198-205 .

[8] L. R. Rabiner, A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proceedings of the IEEE*, 77(2), 1989, pp. 257 - 286 .

[9] A. J. Viterbi, Error bounds for convolutional codes and an asymptotically optimal decoding algorithm, *IEEE Trans. Informat. Theory*, IT-13, 1967, pp. 260 - 269.